

Social signal processing for telemonitoring.

Mark Hartswood, Maria Wolters and Jenny Ure.

'Social data analysis' has emerged as an approach to the collaborative interpretations of data sets using crowdsourcing techniques. It works by breaking an intellectual task into smaller components that can be executed by suitably recruited 'crowds' to create useful, accurate and diverse interpretations of large data sets beyond the capacity of individual experts to achieve. We propose a variant of this approach called Social Signal Processing (SSP) to solve the problem of accurately interpreting increasing volumes of signal data generated by a range of telemonitoring applications.

Social Signal Processing aims to:

- Addresses sense-making problems occurring when data travels between different interpretative contexts (e.g. home and clinic) (Hartswood et al, 2012).
- Aims to blend different expertise that is otherwise separated by contextual and professional boundaries to enable data to be useful in non-local contexts.
- Tackles an increasingly important class of problems arising from a deluge of data (Gershon, 2002) generated by (for example) medical sensors, sensor networks and scientific data streams.

The sense-making problems arise from the fact that data is identifiable (i.e., we know what was measured and when) but obscure (i.e., we don't know what it means because we lack the appropriate interpretive context.) (Hertzog et al, 2010). By using social signal processing, we essentially apply social translucence (Erickson and Kellogg, 2000) to lift the veil of obscurity. However, are we sure that patients want their data to be elucidated?

As Hartzog et al have shown, obscurity is a big part of privacy. Information privacy is at the core of legal and social theories of what privacy is and how privacy is enacted.

"Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others." (Westin 1967, p. 7, cited after Margulis 2011).

Approaches like Petronio's (2002) Communication Privacy Management theory formalise the limits on the flow of information that are imposed by claims to privacy.

- linkage rules determine who patients disclose information to
- permeability rules determine how much others are allowed to know about the information disclosed by the patient
- ownership rules determine how much control the people to whom patients disclose information have about this information.

Such considerations are mostly dealt with under the umbrella of "confidentiality". In our approach to social signal processing, we seek to make them explicit, to ensure that patients and health care professionals can negotiate what is disclosed.

Case study: Telemonitoring

Raw data from telemonitoring solutions is often hard to interpret by remote clinicians who lack firsthand experience of its creation. Recent studies highlight how difficult it can be to establish baseline or 'normal' readings for individual patients (Elwin et al, 2011) and how worrying 'scores' need to be verified by

contacting the patient or carers to weed out false positives¹ (Telescot, 2009) and prevent false negatives² (Anderson et al, 2010). It is hard to mediate a connection between the patient's physiological condition and the physician's diagnostic skill using only remote sensors and analytic software because processes of learning and contextualisation are needed before data can be meaningfully interpreted (Anderson et al, 2010). This poses a dilemma for telemonitoring because the 'friction' (Edwards et al, 2011) apparent when moving data between contexts necessitate expensive workarounds compromising its ability both to deliver cost savings and to provide a scalable solution.

While these basic issues remain unresolved, sensor technologies are making significant advances. Wearable, self-powered sensors able to form ad-hoc networks significantly enhance the potential for gathering continuous streams of diagnostic and research data (e.g. Barnes et al, 2010). Although offering the possibility of pre-emptively diagnosing rare events (e.g. cardiac events), and richer sources of data for stratifying patients into treatment groups, significant problems are posed by the interpretation and management of the high volumes of data generated. Problems of context are magnified when attempting to assign trends and patterns to either environmental or physiological causes. What is more, for each of these potential causes, we need to determine the trade-off between preserving obscurity and therefore privacy and allowing social translucence. High data volumes threaten to overwhelm the expertise available to create interpretations leading to medico-legal hazards when overlooked illness pre-cursors present but buried in a data stream. On the other hand, the private data that would allow creating interpretations may be difficult to gather in real time, attach to the data stream, and encode in a way that respects the patient's information sharing boundaries.

A social signal processing solution acknowledges that interpretation of physiological data has social components and seeks from the outset to build these in to technology solutions in a scalable and integrated way, for example, by enabling the patient, their family and carers to provide a parallel stream of contextual detail that assists accurate interpretation of the telemonitoring data. Challenges for a social signal processing approach include:

- *Facilitating collaborative production and capture of contextual detail.* Social Data Analysis research explores methods both for motivating participation and eliciting relevant and useful crowd contributions. Strategies to maximise participation include delivering benefits such as engrossing gameplay, enhanced status and financial reward. Strategies to improve contribution quality include prompts, exemplars, iterative refinement and rating tasks (Willet et al, 2012). We need to explore how well these strategies transfer into a Social Signal Processing context where motivation and quality become entangled with the divergent priorities and bodies of expertise of professional and lay contributors. The ethical issue here is that maximising participation must not lead patients to unknowingly disclose more than they would be comfortable with – this is an issue with all of the big social media networks such as Facebook.
- *Ensuring the validity of participant created data.* The emerging paradigm within social geography of 'participant sensing' is closely allied with the Social Signal Processing concept and has explored a series of methods for establishing data validity, including triangulation, applying 'common sense' rules, examining data for internal consistency, establishing the reputation of the contributor etc (Burke et al, 2006). In addition, we need to examine the sensitivity of the data, the risk of a breach of privacy, and the implications for other people in the patient's social network of providing that data. Again, we need to explore how these mechanisms translate to Social Signal Processing context, particularly in settings where access to resources can be at stake raising the prospect of 'gaming' behaviour.

¹ Even with an established baseline, the data can be complicated by local factors that are independent of the disease state, for example, recent exertion or stresses specific to the patient [2].

² Anderson et al report false negatives when monitoring data is interpreted without the involvement of the patient [4].

- *Exploring the role of analytics and creating an appropriate composite of computational and social interpretations of the data.* The Social Signal Processing perspective recognises that interpretation of sensor and other data streams is always a blend of social and computational components. This contrasts strongly with 'pervasive computing' approaches that seek to eliminate the human contribution by creating increasingly elaborate monitoring regimes, rule sets and algorithms to entirely automate capture of context (e.g. Copetti et al, 2009). We believe that hybrid approaches where people supply missing contextual detail is by far the more effective, efficient and achievable approach. The challenge for SSP is to produce engineering strategies that can orchestrate working configurations of analytics, instrumentation (e.g. via sensors or sensor nets) and human contributions that acknowledge the patient's privacy and do not treat them as an entity to be observed, but as somebody who has control over the information that is distributed.
- *Blending sensor data and contextual detail into appropriate sharable visualizations.* This concerns the appropriate forms of the shared analytic objects created by a SSP approach. These should be capable of being interpreted and elaborated by diverse stakeholders, support negotiation and iterative refinement of contributions. It is here that the multidisciplinary nature of SSP research comes to the fore with the requirements to understand and apply shared sense-making, human-factors, data analytics and visualisations in a distributed collaborative setting. Again, these visualisations need to be detailed enough to allow detection of disease relevant patterns, but obscure enough to protect individual detail.
- *Designing to support emergent individual and collaborative interpretative practices.* Pilots reveal both how existing responsibilities are redistributed between stakeholders and how new forms of work emerge on the introduction of new telemedicine solutions. It is common for the patient to become engaged in diagnostic practices and for auxiliary healthcare workers to play an 'unscripted' role in mediating interpretations of 'signal' data (Mort et al, 2003; Oudshorn, 2008). Thus a SSP approach should proceed on the assumption that interpretative practices will become distributed through the system rather than being the province of individual agents. Instead of resisting or ignoring these processes, SSP aims to encourage and exploit distributed interpretive effort by providing appropriate support and resources for each individual's contribution. When engaging with patients in this collaborative sensemaking activity, there also needs to be a component dedicated to establishing boundaries. This is hard – the transparency paradox (c.f. Nissenbaum 2011) shows that if all ways in which information could be used is made transparent, it is often unintelligible to the person whose consent is being sought.
- *Supporting reconfiguration of roles within a telemonitoring solution.* Telemedicine solutions have the potential to invert the usual roles of patients and health care providers. Given access to their own data patients can take a stronger role in managing their own condition. Patients pooling data can create community-led interpretation or motivational mechanisms and embark upon community driven research. Whilst of huge potential benefit, this risks marginalising the 'medical voice', and opens up a series of confidentiality issues that turn upon community standards rather than fiduciary obligations. Furthermore, issues of translucence and obscurity have the potential to cut both ways. Not only does telemonitoring reveal more about the patient, it also shines a stronger light on the clinical decision-making process by making more explicit the data and the decision mechanics. Without careful attention, health care professionals may prefer system configurations that trade clinical effectiveness for medico-legal safety.
- *Engineering platforms and services that enable SSP applications to rapidly assembled for new domains.* Although our work is predicated on a case study in telemedicine, we will identify stable and portable approaches that can be abstracted into reusable components, services and templates to facilitate engineering an SSP across a spectrum of telemedicine applications.

- Anderson T, Pernille B, Kensing F and Moll J (2010) Designing for collaborative interpretation in telemonitoring: Re-introducing patients as diagnostic agents. *International Journal of Medical Informatics*. 80(8) 112-126.
- Barnes M, Conway C, Mathews J and Arvind D K, "ENS: An Energy Harvesting Wireless Sensor Network Platform". The Fifth International Conference on Systems and Networks Communications, ICSNC 2010
- Burke J, Estrin D, Hansen M, Parker A, Ramanathan S, Reddy S and Srivastava M B (2006) Participatory sensing. UC Los Angeles: Center for Embedded Network Sensing. Retrieved from: <http://escholarship.org/uc/item/19h777qd>
- Copetti A, Loques O, Leite J C B, Barbosa T P C and da Nobrega A C L (2009) Intelligent Context-Aware Monitoring of Hypertensive Patients. 3rd International Conference on Pervasive Health. 1- 3 April. London.
- Edwards P, Mayernik M S, Batcheller A, Bowker G and Borgman C (2011) Science Friction: Data, metadata and collaboration. *Social Studies of Science*. 41(5) 667-690.
- Elwin G, Hardisty A R, Peirce S C, May C, Evans R, Robinson D K R, Bolton C E, Yousef Z, Conley C, Rana O F, Gray W A and Preece A D (2011) Detecting deterioration in patients with chronic disease using telemonitoring: navigating the 'trough of disillusionment' 18, 896–903
- Erickson, T., & Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction*, 7(1), 59-83. doi:10.1145/344949.345004
- Gershon D (2002) Dealing with the data deluge. *Nature*. 416, 889-891.
- Hartwood M, Procter R, Taylor P, Blot L, Anderson S, Rouncefield M and Slack R. Problems of data mobility and reuse in the provision of computer-based training for screening mammography. In: CHI '12 Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems. New York: ACM. 2012. p. 909-918.
- Hartzog, Woodrow and Stutzman, Frederic D., The Case for Online Obscurity (April 29, 2010). *California Law Review*, Vol. 101. Available at SSRN: <http://ssrn.com/abstract=1597745> or <http://dx.doi.org/10.2139/ssrn.1597745>
- Mort M, May C R and Williams T (2003) Remote doctors and absent patients: Acting at a distance in telemedicine. *Science, Technology and Human Values* 28(2) 274-295.
- Margulis, S. T. (2011). Three Theories of Privacy. In S. Trepte & L. Reinecke (Eds.), *Privacy Online* (pp. 9-17). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-21521-6
- Nissenbaum, H. (2011). A Contextual Approach to Privacy Online. *Daedalus*, 140(4).
- Oudshoorn N (2008) Diagnosis at a distance: the invisible work of patients and healthcare professionals in cardiac telemonitoring technology. *Sociology of Health and Illness*. 30(2) 272-288.
- Petronio, Sandra (2002): *Boundaries of privacy: Dialectics of disclosure*. Albany, NY: SUNY Press.
- Telescot (2009) Reconfiguring Care for Chronic Disease with Home Monitoring: a pilot study report on telemetry-supported care for COPD in the Lothian Region. (<http://www.telescot.org/uploads/4/5/9/4/4594120/telescot-copd-pilot-report.pdf>, accessed: 4/9/2012)
- Westin, Alan (1967): *Privacy and Freedom*. New York: Atheneum
- Willet W, Heer J and Agrawala M (2012) Strategies for Crowdsourcing Social Data Analysis. CHI'12 , May 5–10, 2012, Austin, Texas, USA.